DOCUMENT RESUME

ED 331 894 TM 016 661

AUTHOR Neill, Monty
TITLE Why the United States Does Not Need a National Test: Testimony to the House Subcommittee on Select Education.
INSTITUTION National Center for Fair and Open Testing (FairTest), Cambridge, MA.
PUB DATE 23 Apr 91
NOTE 22p.
PUB TYPE Legal/Legislative/Regulatory Materials (090) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Change; Educational Objectives; Elementary Secondary Education; Equal Education; *National Programs; School Based Management; *Testing Programs
IDENTIFIERS FairTest; *National Assessment of Educational Progress; Performance Based Evaluation

ABSTRACT

Based on the FairTest perspective, this paper argues that the United States does not need a national test to measure progress toward the nation's educational goals and that such a test would have adverse impacts on low-income and minority students. The National Assessment of Educational Progress (NAEP) should remain an indicator system but should use more performance-based methods in its assessment. National testing proposals are usually based on the false premise that measurement itself will produce positive change in education. A national examination could undermine needed and emerging reforms such as school-based management and shared decision making. A national test would tend to centralize decision making, making education less accountable to parents, students, teachers, and the community. A national examination would not promote educational equity. The weaknesses of multiple-choice examinations are also dangers inherent in a national examination. Recommendations are made for appropriate educational reform; these include development and implementation of performance-based assessment methods. Attachment A is a statement on proposals for a national test, which summarizes the reasons for opposing a national test. Attachment B is an open letter, which discusses 10 concerns and recommendations about the roles of the NAEP and the National Assessment Governing Board. The names of signers of both attachments are listed. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

FairTest

National Center for Fair & Open Testing

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Monty Neill
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED331894

Why the United States Does Not Need a National Test:

Testimony to the House Subcommittee on Select Education

by

Monty Neill, Ed.D., Associate Director,

National Center for Fair & Open Testing (FairTest), Cambridge, Mass.

April 23, 1991

Ladies and Gentlemen of the House:

Thank you very much for inviting FairTest to appear at this important hearing.

Based on an examination of existing proposals, FairTest concludes that most current efforts to establish a national test to measure progress toward the nation's educational goals will hurt, not help, our nation's efforts to improve school quality. The damage will fall most heavily on low-income and minority-group students. We therefore urge the House of Representatives to support education reform by ***not*** implementing a national exam at this time. The House should, however, support efforts to introduce new assessment methods as part of implementing school reform.

The House also should not turn the National Assessment of Educational Progress (NAEP) into a national examination by allowing comparisons below the state level or the use of NAEP tests or items for district or state use. NAEP should remain an indicator system, but should use more performance-based methods in its assessments.

To address these two points, I will first discuss the reasons why a national examination should not be implemented at this time, with particular reference to the harm such an examination would cause to low-income and minority group students. Secondly, I will discuss proposed expansion of NAEP.

National testing proposals largely are based on the false premise that measurement by itself will produce positive change. Recent history shows this is not true: During the 1980s, U.S. school children became probably the most over-tested students in the world -- but most of the desired educational improvements did not occur.[1] FairTest research, reported in *Fallout from the Testing Explosion*, indicates that our schools now give more than 200 million standardized exams each year and the typical student must take several dozen before graduating.[2] Adding more testing will no more improve education than taking the temperature of a patient more often will reduce his or her fever.

In contrast, successful educational reform must include restructuring curriculum, instruction, textbooks and other materials, school governance, and teacher education, as well as assessment. This must be done for ***all*** students. What we need to create are schools as communities of and for learning.

TM016661

BEST COPY AVAILABLE

342 Broadway, Cambridge, Mass. 02139 (617) 864-4810 FAX (617) 497-2224

To move toward that goal, teachers, administrators, other school personnel, parents, students, community members, and government must all be involved in an open and democratic process of defining our educational goals -- at the local, state and national levels -- so that we can agree, for example, on what it means for all students to be competent in different areas. On that basis, we can then determine how to make the changes required to reach the goals, including a decision on whether to institute a national test. Most current proposals for a national test, however, seek to test *before* necessary decisions about the goals of school reform have been made. This likely will lead to the backdoor imposition of a national curriculum, without public discussion.

Indeed, having a single national test raises the issue of the control of education. If the test becomes important, as all testing proponents want, those who control the test could control curriculum and instruction, particularly if decisions about curriculum and instruction have not been arrived at before the test is constructed, and maybe even if those decisions have been reached.

A national exam should not be allowed to undermine such needed and emerging reforms as school-based management and shared decision-making. By centralizing decision-making, centralized national testing most likely will make education less, not more, accountable to parents, students, teachers and the community. If the test is centrally controlled, to whom could parents, teachers and communities appeal if they disapprove of the curricular decisions and instructional methods imposed through the test?

A second fallacy underlying proposals for a national test is that the US needs a national exam because its students do not perform as well on international comparisons and because the US educational system must improve to enable economic competitiveness. FairTest supports educational improvement, though the reason should not be reduced to economics. But educational improvement does not require a national exam. Neither Germany nor Japan has a national examination akin to any of the proposals that have been made in this country. Germany also does not have a national curriculum. If these nations provide a better education to their children, it cannot be because they have a national examination.[3]

A third major problem that has not been addressed in any national testing proposal is the question of equity. No one test should become a national gatekeeper that perpetuates our nation's unfortunate history of unfairly sorting students by race and class.

A national test could end up being used to determine high school graduation, employment, and entrance into higher education. Due to unavoidable measurement error and bias, many students who fail a test will, in reality, be as capable as many who pass. Research indicates that those who fail but should have passed will be disproportionately from low-income and minority-group backgrounds.[4] FairTest agrees with the National Commission on Testing and Public Policy -- a Ford Foundation-supported body that studied testing for three years -- that, because of the bias and error, no one test should ever be the sole or primary basis for making an important educational decision.[5]

No new exam or examination system should be implemented without assurances -- in practice, not rhetoric -- that all students will be given an equitable opportunity to pass the tests. It should also be clear how the tests will be used to improve education, not just continue to sort students, before any national test is implemented.

Dangers of Multiple-Choice Testing

FairTest recognizes that there are two different types of proposals for national testing. One type will rely essentially on multiple-choice testing; the other calls for performance-based assessment. These two approaches are quite different. They are the difference between testing *what students should know* and *what students know how to do*.[6]

The first approach quickly leads to multiple-choice testing of arbitrary facts and isolated skills, unconnected to the way knowledge is used in the world. Multiple-choice and short answer tests cannot adequately assess problem-solving or the ability to create and use knowledge.[7] Higher order thinking requires the student to define the problem, to consider and attempt various solutions to problems which are ill-structured and may have more than one correct solution, and to produce knowledge, not merely recognize answers.

Because multiple-choice/short answer testing cannot directly assess higher order capabilities, a test comprised of such items will not inform us as to the problem-solving and knowledge-creating capabilities of our students. We know from research, however, that student abilities in these areas are very limited. This has been caused largely because of schools' failure to teach them in any subject area to more than a few students. Even the best high school students typically do not know how to problem-solve using the approaches and methods a professional uses.[8] Yet research also shows that problem-solving, knowledge-creating approaches can be used even with very young children.[9]

If a test is important -- as a national test is sure to be -- then teachers will teach to it.[10] Because multiple-choice tests cannot directly measure higher-order skills, teaching to the test reduces or eliminates instructional time spent on the higher skills. Instruction is reduced to drilling for multiple-choice exams and the curriculum is reduced to the test. Multiple-choice testing precludes a curriculum based on thinking, investigating, problem-solving and using creativity, because the test cannot measure those things.

Any norm-referenced test must make cultural assumptions through the language used and the experiences the test treats as normal or common. The tests have assumed that the test-taker comes from a white middle-to-upper class background. Students who come from a different culture due to class, race or language factors are automatically at a disadvantage on most tests. For technical reasons in constructing a norm-referenced test, items that a minority test-taker is likely to get correct but a white middle-class test-taker is likely to get wrong are excluded from the tests.

The major initial use of tests in schools was to sort students, and this remains a primary use of the tests, often starting at a young age. In Boston, Massachusetts, for example, the grade two, multiple-choice reading test is used to determine entry into advanced work classes; in turn, these classes largely determine who will be able to enter the city's examination schools; and while most examination-school graduates attend college, very few other Boston high school graduates enroll in college.[11] In short, a multiple-choice grade two reading test that assumes a middle-class cultural background largely determines the educational opportunities of Boston's youth.

Low-income and minority group children are disproportionately tracked into low-level courses, often on the basis of test scores. In these courses they are typically subjected to routine, basic-skills "drill and kill" oriented toward increasing test scores on multiple-choice

tests. As a result, these students are least exposed to higher-order academic skills. In open-ended, performance-based assessments in Massachusetts, for example, children from low-income areas scored very low, and investigation showed these children had not been taught the abilities being assessed.[12]

Multiple-choice tests perpetuate the false idea that first students learn basic skills, then they learn higher skills. Cognitive psychological research has demonstrated that learning involves active thinking and to enhance learning the student must be actively engaged.[13] Test-driven schools produce higher test scores, but not students who are able to think. This problem affects low income and minority-group children most heavily.

A predominantly multiple-choice test may include a writing sample. A typical short writing sample requires a student to write several hundred words on a topic he or she may or may not know anything about and may or may not care about, in a short period of time, with no chance for research, discussion (that is called cheating), or serious revision, for no purpose except the test. This is the sort of writing assessed by NAEP. If the purpose of writing is to communicate, then a typical test writing sample cannot legitimately be called writing at all. As with multiple-choice testing, it sends the wrong message about the goals of education.

Short writing samples also may underestimate the ability of students rated as low performers. For example, timed writing samples do not allow time for revision, which may particularly harm students whose first language is not English. In a study of portfolio writing in Durham, NH, researchers found that students who scored low on writing tests tended to raise their performance level to the middle range on portfolios where they had extended writing time and could write on more meaningful topics.[14]

Multiple-choice and short-answer tests are not very useful to teachers or policymakers. The reason, in both cases, is that the test results do not help the teacher or policymaker decide what to do next. If Johnny cannot multiply, the test cannot explain why. If Maria's whole class cannot multiply, the test does not provide information on what should be done.

What standardized multiple-choice tests do best is help sort students. It is what they were invented to do. But if we are serious about reforming education so that all students can learn the things we deem important, then we must stop relying on tests that have as their only real use the sorting of students.

In sum, implementing a national multiple-choice exam will mislead the public about the nature of the problem and the requirements of real change, block positive school reform (including the use of new methods of assessment), hinder students' ability to develop the kinds of intellectual competencies they need to develop, probably perpetuate sorting students by class and race, and ultimately undermine public education.

No proposal that suggests using more than a small proportion of multiple-choice items in a national examination should be given any serious consideration by the House. At most, multiple-choice could be used as part of a sampling program to gather limited information about student acquisition of a narrow range of knowledge. There is no reason to test every student for this purpose, and such a purpose should never be allowed to dominate education, as it too-often now does.

Because of current technical limitations, any proposal to assess our nation's students inexpensively and in the near future will, of necessity, be a multiple choice test. An example is the proposal by Educate America to test all high school seniors in six subjects for $30 -

$50 each. They claim their tests would be "state of the art" and include performance-based components. But the SAT, which is entirely multiple-choice, costs $16 for just two tests. The Educate America plan would have to be a multiple-choice test. Such proposals must be rejected.

Performance-Based Assessment

By contrast, students should be assessed on what they know how to do. To know how to do something includes knowing factual content. This method of assessment corresponds to how people learn. They learn by integrating new information or experiences into the intellectual frameworks they already possess, which in turn enables them to refine and improve the frameworks.

Assessing what students know how to do is based on students' doing real work. There are many ways for students to demonstrate intellectual competence in and across the subject areas. Performance-based assessments can be based on regular student classroom work - projects, research, writings, products, self-reflection, teacher evaluation, exhibitions, and performances - that can be organized and summarized in portfolios. In turn, the portfolios can be examined by outside people -- teachers, other parents, trained examiners -- to determine the quality of the portfolios and the kinds of work students are doing. Vermont for example, is working on this method.

Performance-based assessments can also be examinations administered from outside the classroom. These can include open-ended, complex problems requiring the student to figure out what to do, solve the problems, and explain what he or she did. Or they can be exhibitions, performances and products, such as now done in science fairs, Scout Merit Badges, Advanced Placement Art, and many performing and applied arts. These often can be exams that are worth teaching to, unlike multiple-choice tests. Arizona, California, Connecticut and Maryland are among the states implementing these types of exams.

Taken together, in-class and externally-developed performance-based exams can encourage real work, model high standards, spur improvements in teaching and curriculum, produce instructionally useful information for teachers and students, and provide information based on real activities about student progress. Assessment can play an important part in developing communities of and for learning.

Cautions on a National Performance-Based Examination System

However, support for performance-based assessments does not mean such assessment should immediately be transformed into a national examination system, such as that proposed by the Learning Research Development Center and the National Center for Education and the Economy (LRDC/NCEE).[15] There are many reasons why this is the case. Among them are:

-- We have not yet completed the process of discussing and debating what we want our educational systems to be. Many complex issues of educational reform, involving curricular goals and standards, instructional methods, assessment methods, school structure and governance, and collection of information, largely must be resolved *before* the question of whether a national examination system is desirable can be answered. To do otherwise is to put the cart before the horse.

-- In general, the proposal does not adequately address equity issues that must be solved for the system to be fair. Changing assessment will not by itself reduce inequities. All students must be assured a fair opportunity to learn how to work within a thinking curriculum that uses performance-based assessments. As in the Massachusetts case noted earlier, this is not now the case. Additionally, the goal of "initial mastery," called for in a number of the proposals, could encourage sorting and tracking students according to who can best or most quickly reach the goal. This danger needs to be seriously addressed to try to ensure structures and processes, including in the realm of assessment, that are inclusive and reduce tracking and other kinds of sorting. Also, because competence in subject areas may best be expressed in languages other than English for some students, the option of being assessed in other languages must be available before any exams are imposed. Finally, while FairTest believes that performance-based assessments can be used fairly and can even assist in overcoming racial and cultural biases or ignorance, it will not happen automatically. Virtually no research has yet been done on this topic. It would be dangerous to implement a national performance-based examination system without building-in methods to ensure fairness and equity.

-- Imposing a national examination will not address the issues of rigid and bureaucratic school governance and structure, low-quality textbooks, and inadequate schools of education. Improving assessment needs to be considered as one part of integrated systemic change.

-- The proposal calls for national boards to set standards. It could create a national school board that, by setting curriculum standards, will lead to a centralized, national education system. Because the consequences of such actions cannot now be known, but may include undermining democratic control of education, we should not rush into that process.

-- Staff development is central to school reform, but is not adequately addressed in the LRDC/NCEE proposal. If teachers are to teach to performance-based assessments, to teach the "thinking curriculum," they have to know how to do so. This involves developing the ability of our nation's 2-1/2 million teachers to teach and assess in new ways. To be effective, school reform must include the active participation of those who will implement the changes. We cannot impose new assessments on teachers, change nothing else, and say "Do it."

-- We simply do not know whether it is feasible to construct a national examination system. The whole process, particularly the calibration, could prove to be too complex, expensive and unwieldy to work. (Calibration is the process by which student results on different exams can be determined to be equivalent to each other and to national standards; it would mean, for example, calibrating one states history exam to another's, even when the precise topics might not be the same or when one state insists on essays but another allows videos or public performances as well as essays. This is typically a labor-intensive process that is also valuable for staff development. However, England recently dropped a moderation process from its national exam process because it was too expensive. Moderation is the process by which teachers help shape standards and learn to grade papers, products and performances uniformly so as to produce consistent and reliable results, and is therefore akin to calibration in many regards. Moderation is valuable and necessary and must be included in

any performance-based system, but doing it on a national level on top of state and local levels may be too much as well as unnecessary for educational improvement.

-- When the complexities and expense of the proposal become clear, the portfolios and projects could end up being reduced to very limited exams. There even could be a return to multiple-choice and short-answer exams. Such a retreat would have destructive curricular effects and undermine all aspects of educational improvement.

-- The proposal is not conceived of as one part of an overall educational information system. Having assessment outcome information on education is not useful unless we also have adequate information on inputs (money, teaching staff, building quality, etc.), processes and programs (curriculum, instructional methods, textbooks and materials, class size, role of tracking, governance and school organizational structure, etc.), and additional outcome data (employment and further education of graduates, dropout rates, etc.). This information should be obtained without harming education -- unlike what has happened with multiple-choice tests. Schools and programs should be evaluated on a comprehensive range of indicators of their quality as communities that support learning for all students.

-- Finally, the money spent just on the nationalizing aspect of the implementation of new assessments might better be spent on supporting comprehensive educational reform rather than on calibrating exams.

Recommendations

There is no one, simple method of putting a national education reform process into motion in the right direction. It is a process that can and is happening at all levels: the classroom, the school, the district, the state, consortia that include all of these, and at the national level. It is not and will not be a smooth and easy process. But as good practice becomes available to replicate, as improved curriculum and assessments become more widely known, as our nation's willingness to improve education for all children continues to grow, then we can expect to see real progress.

FairTest is far from alone in this view. Over two dozen national civil rights and education organizations joined FairTest's Campaign for Genuine Accountability in Education in issuing a "Statement on Proposals for a National Test." These organizations include the NAACP, the Mexican-American Legal Defense and Education Fund, the Puerto Rican Legal Defense and Education Fund, the National Education Association, the National PTA, the National Association for the Education of Young Children, the American Association of School Administrators, and both National Elementary and Secondary School Principals Associations. The statement urged "the Bush Administration and the Congress to support education reform by not implementing a national exam at this time." (The "Statement" and list of signers is appended as "Attachment A.")

The federal government can proceed in one of two ways. It can impose a national test that runs the risk of short-circuiting the process of school reform. Or it can find ways to support school reform activities without imposing a national test.

FairTest concludes that the House of Representatives should not propose a national exam either immediately or to be in place within any fixed timetable, such as five or ten years. Rather, FairTest urges the federal government to take the following steps to improve education and assessment:

-- Assist states and districts, acting in consortia, to develop and implement performance-based methods of assessment.

-- Assist state and district teacher education and staff development programs.

-- Assist the subject area groups, such as those in math, English, social studies and science, to develop and disseminate model curricula, standards and assessments.

-- Re-examine the instances in which the federal government requires standardized multiple-choice testing, particularly for the Chapter I program. The testing requirements virtually force programs into being test-coaching programs, though that, as explained above, is a poor educational method.

-- Consider how assessment information can best be included as one element of school reform activities and one part of an indicator system, and not view assessment in isolation.

In all of these efforts, concern for fairness and equity must be included. Promises and hopes will not suffice; rigorous planning to ensure equity is necessary.

Only after these educational reform processes have been implemented and evaluated over a period of time should the federal government consider whether it is desirable or feasible to link the newly developed local and state performance-based assessments to each other and to national standards or curricular frameworks.

The National Assessment of Educational Progress

The National Assessment of Educational Progress should remain as a national indicator. To turn it into some kind of a national test will end up destroying its current usefulness and will produce the drawbacks discussed above. In particular, NAEP should not be used below the level of state-level comparisons. FairTest doubts that state-level comparisons will be of real use to educators and urges that state comparisons not be approved beyond trial measures unless experience and research demonstrate how the comparisons will be used to improve education. NAEP should, however, include far more performance-based assessments and provide technical assistance to districts, states and consortia who are implementing performance-based assessment.

The National Assessment Governing Board (NAGB), has proposed substantial expansion of NAEP, to include use of NAEP items down to the individual level. If this occurs, teachers will begin to teach to NAEP, producing the Lake Wobegon effect.[16] (The "Lake Wobegon Effect," named after Garrison Keilor's mythical town where "all the children are above average," describes the inflated and misleading test scores that come from teaching the test.) This will eliminate the possibility of using NAEP as an indicator and the nation will no longer be able to rely on the accuracy of NAEP data.[17]

Last spring, in response to the NAEP expansion proposals, FairTest asked the organizations supporting its Campaign for Genuine Accountability in Education to endorse an "Open Letter to Congress, Bush Administration, the Governors on NAGB and NAEP expansion." The statement was endorsed by dozens of national and local education and civil rights organizations and prominent individuals. The statement detailed the problems with the expansion proposal. I attach a copy of the statement and the list of signers as part of this testimony ("Attachment B").

Conclusion

Let us be clear. FairTest is not arguing against accountability or for slowing down school reform. Nor is the issue one of the need for "standards." Rather, the central issue is how we define education. We are saying that we need school reform, not more testing. We need genuine accountability, not test scores from multiple-choice or short-answer exams, and we don't need to jump aboard an examination train heading into trackless terrain.

Our nation must not be misled into thinking more testing will solve our educational problems. Instead, we must construct plans for reform that include assessments which can be used to help student learning, guide educational improvement, provide information for accountability, and assist the goal of equity, but not block progress or harm students. Our nation will be far better served to take the time to do the job well, than to act hastily and poorly with destructive results.

Endnotes

1. National Assessment of Educational Progress Reports on Reading (1990), Math (1988) and Writing (1990) (Princeton, NJ: Educational Testing Service). Looking at the data for the 1980's, a decade of massive increases in testing, we see that reading scores declined slightly for the youngest group, were flat for the middle group, and only increased slightly for older children. The math performance was not as bad, showing some gains at the lower skill levels, but not at higher skill levels; and there were no gains in writing results. Minority-group children did improve at the lower-skill levels, closing the gap with white children, but these gains were erratic. Black 9-year-olds made no gains in the 1980's in reading, and gains were slight or non-existent for 9-year-old black and Hispanic children in math. Black and Hispanic children generally failed to improve at the higher-skill levels.

2. Medina, Noe and D. Monty Neill, *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools* (Cambridge, MA: FairTest, 3rd Ed., 1990).

3. Kelley, E.W. *Can National Tests Affect the Quality of Education?* Testimony before the Subcommittee on Elementary and Secondary Education of the House Committee on Education and Labor, March 14, 1991. I.C. Rotberg, "I Never Promised You First Place," *Phi Delta Kappan* (December 1990) pp. 296-303, and Kelley indicate that the commonly cited rankings from international comparisons are quite flawed; both, however, do think areas of US education need improvement.

4. Hartigan, John A. and A.K. Wigdor. *Fairness in Employment Testing* (Washington, DC: National Academy Press, 1989) explains statistically why this is the case even in the absence of bias in exams. For a bibliography on bias, see Medina and Neill, op. cit.

5. National Commission on Testing and Public Policy, *From Gatekeeper to Gateway: Transforming Testing in America* (Chestnut Hill, MA: Author, 1990);

6. Susan Harman, drawing particularly on the work of people associated with the Coalition for Essential Education, uses this formulation very clearly in "National Tests, National Standards, National Curriculum," *Language Arts* (January 1991: 49-50).

7. For a general critique of multiple choice testing, see Neill, D. Monty and Noe J. Medina, "Standardized Testing: Harmful to Educational Health," *Phi Delta Kappan* (May 1989: pp. 688-697), and Medina and Neill, *op. cit.* For analysis of why multiple-choice testing cannot assess higher order thinking, c.f. Fredericksen, Norman, "The Real Test Bias: Influences of Testing on Teaching and Learning," *American Psychologist* (March 1984: pp. 193-202); National Commission on Testing and Public Policy, *op. cit.;* Resnick, Lauren B. and Daniel Resnick, "Assessing the Thinking Curriculum: New Tools for Educational Reform," in B.R.

Gifford and M.C. O'Connor, eds., *Future Assessments: Changing Views of Aptitude, Achievement, and Instruction* (Boston: Kluwer Academic, 1989).

8. "Beyond the Bubble," a mini-conference at the April 1990 national conference of the American Educational Research Association, Boston, MA.

9. Kamii, Constance and Mieko Kamii, "Why Achievement Testing Should Stop," and Engel, Brenda S., "An Approach to Assessment in Early Literacy," in Kamii, Constance, ed., *Achievement Testing in the Early Grades: The Games Grown-Ups Play* (Washington, DC: National Association for the Education of Young Children, 1990).

10. Madaus, George. "The Influence of Testing on the Curriculum." *87th Yearbook of the National Society for the Study of Education, Part I: Critical Issues in the Curriculum* (1988: pp. 83-121). Also, Resnick and Resnick, *op. cit.*

11. Massachusetts Advocacy Center. *Locked In/Locked Out* (Boston: Author, 1990).

12. Reported by Elizabeth Badger of the Massachusetts Department of Education to the Alternative Assessment Conference of the Technical Education Research Centers of Cambridge, Mass., March 8-10, 1991.

13. Resnick and Resnick, *op.cit.*

14. Simmons, J. "Portfolios as Large-Scale Assessment," *Language Arts* (March 1990) pp. 262-268.

15. Learning Research and Development Center and National Center on Education and the Economy. *Setting a New Standard: Toward an Examination System for the United States, A Proposal* (Pittsburgh and Rochester: Author, October 1990) is the most comprehensive of the reports and documents proposing a national performance-based examination system.

16. Named after Garrison Keilor's "Lake Wobegon" where all the children are above average. Research has shown that this effect is true -- more than half the students, districts and states are "above average" -- and that teaching to the test is the primary cause of the effect. See: L.A. Shepard, *"Inflated Test Score Gains": Is It Old Norms, or Teaching the Test?* (Los Angeles: UCLA Center for the Study of Evaluation, CSE Technical Report 307, 1990).

17. Daniel Koretz from Rand and Robert Linn from the University of Colorado both testified strongly to this effect before the House Subcommittee on Elementary and Secondary Education on March 13, 1991.

Campaign for Genuine Accountability in Education

Statement on Proposals for a National Test

As educators, parents, and civil rights advocates, we strongly support improving assessment as part of school reform. However, we believe that most current efforts to establish a national test to measure progress toward the nation's educational goals, such as the proposal from Educate America, will hurt, not help, school quality.

We therefore urge the Bush Administration and the Congress to support education reform by not implementing a national exam at this time. Rather, they should support efforts to introduce new assessments as part of implementing school reform and genuine accountability.

Successful educational reform must include restructuring curriculum, instruction, school governance, and assessment. This includes developing the ability of our nation's 2-1/2 million teachers to teach - and assess - in new ways. Teachers, administrators, other school personnel, parents, students, community members, and government must all be involved in an open and democratic process of defining our nation's educational goals so that we can agree, for example, on what it means for all students to be competent in different areas. On that basis, we can then determine how to make the changes required to reach the goals, including a decision on how best to assess progress toward the goals. Most current proposals for a national test, however, seek to test before necessary decisions about the goals of school reform have been made. This likely will lead to imposition of a national curriculum without public discussion that will block our nation's progress toward high-quality education for all.

Most current proposals call for creation of a low-cost test that will be administered to all students in the near future. Such proposals suffer from several fatal flaws. First, they assume that measurement by itself will produce positive change. Recent history shows this is not true: During the 1980s, U.S. school children became the most over-tested students in the world -- but the desired improvements did not occur. Our schools now give more than 200 million standardized exams each year and the typical student must take several dozen before graduating. Adding more testing is clearly not the way to improve education any more than taking the temperature of a patient more often will reduce his or her fever.

Second, because of cost and time factors, such a test inevitably will be mostly multiple-choice. Because teachers will be pressured to teach to the test, schooling will be reduced even more to test-coaching that will not include learning to think and create and use knowledge in real-world settings. Implementation of such exams therefore will mislead the public about the nature of the problem and the requirements of real change, block positive school reform (including the use of new methods of assessment), hinder students' ability to develop the kinds of intellectual competencies they need to develop, and ultimately undermine public education.

Instead of implementing a national exam at this time, we urge Congress and the Administration to take the following steps to improve education and assessment:

- Assist states and districts, acting in consortia, to develop and implement performance-based methods of assessment.

- Assist state and district teacher education and staff development programs.

- Assist the subject area groups, such as those in math, English, social studies and science, to develop and disseminate model curricula, standards and assessments.

- Ensure that the National Assessment of Educational Progress (NAEP) remains a national monitoring system and focuses on developing high-quality, performance-based assessments; not consider expansion of state comparisons under NAEP until adequate research and discussion on the effects has been completed; and continue the prohibition on comparisons below the state level unless and until NAEP exams are revised to meet the criteria of being performance-based, based on national standards reached by public consensus, and able to be used without undermining NAEP's role of a national indicator that uses matrix sampling.

Only after these educational reform processes have been implemented and evaluated should the Congress and the Administration consider whether it is desirable or feasible to link the newly developed local and state performance-based assessments to each other and to national standards or curricular frameworks.

We are not arguing against accountability or for slowing down school reform. To the contrary, we are saying that we need school reform, not more testing. We need genuine accountability, not test scores from multiple-choice or short-answer exams.

Our nation must not be misled into thinking more testing will solve our educational problems. Instead, we must construct plans for reform that include assessments which can be used to help student learning, guide educational improvement, provide information for accountability, and assist the goal of equity, but not block progress or harm students. Our nation will be far better served to take the time to do the job well, than to act hastily and poorly with destructive results.

List of Signers

Advocates for Children of New York, Inc.
American Association of School Administrators
APPLE Corps, Inc.
ASPIRA Association
Atlantic Center for Research in Education, North Carolina
Leonard Beckum, Duke University*
California Tomorrow
Center for Women Policy Studies
Central Park East Secondary School, New York

Sandra Cox, Testing Committee Chair, Association of Black Psychologists*
Dr. Alfredo de los Santos, Maricopa Community Colleges, Arizona*
Harold Dent, Psychological & Human Resources Consultants, Inc.*
Designs for Change, Chicago, Illinois
Education Law Center, New Jersey
Dr. Pamela George, North Carolina Central University*
Girls Incorporated, Washington, DC
Leslie Hart, Brain-Compatible Education Information*
Asa Hilliard III, Georgia State University*
Institute for Learning and Teaching, Minnesota
Intercultural Development Research Association, Texas
Kentucky Youth Advocates, Inc.
Massachusetts Advocacy Center
META Inc.
Susan Metz, Human Resources Academy, New York*
Mexican American Legal Defense and Education Fund
Mississippi Human Services Agenda
National Association for the Advancement of Colored People
National Association for the Education of Young Children
National Association of Elementary School Principals
National Association of Secondary School Principals
National Center for Fair & Open Testing (FairTest)
National Coalition of Title I Chapter I Parents
National Education Association
National PTA
Fred Newmann, University of Wisconsin*
Organization of Chinese American Women
Panasonic Foundation
Vito Perrone, Harvard Graduate School of Education*
Representative C.J. Prentiss, Ohio State Legislature*
Project on Equal Education Rights, New York
Puerto Rican Legal Defense and Education Fund
Rethinking Schools, Wisconsin
William Robinson, District of Columbia School of Law*
William Schipper, National Association of State Directors of Special Education*
Susannal Sheffer, Growing Without Schooling, Massachusetts*
Southern Association on Children Under Six
Southern Christian Leadership Conference
Southern Regional Council, Inc.
Chuck Stone, University of Delaware*
Student Advocacy Center, Michigan
Representative Vernon Sykes, Ohio State Legislature*
Sara Wallace, National Council for Social Studies*
Whole Language Umbrella

*Organizations listed for identification purposes only

FairTest
National Center for Fair & Open Testing

Testimony of Monty Neill: Attachment B

June 15, 1990

OPEN LETTER TO CONGRESS, BUSH ADMINISTRATION, THE GOVERNORS ON NAGB AND NAEP EXPANSION

Over the past several months, the National Assessment Governing Board (NAGB) has taken several actions which, considered together, raise serious concern over the future direction of the National Assessment of Educational Progress (NAEP). As a group of education and civil rights organizations active in school reform issues, we are addressing our concerns to Congress, the Administration and the National Governors Association so that all responsible parties understand the nature of these problems and carefully monitor developments in NAGB and NAEP. It is important to note that we are not writing to oppose the national assessment, but to help ensure that it plays a constructive, not harmful, role in reforming our nation's educational systems.

The actions of the Governing Board, taken together, go far beyond the level of activity authorized in the National Assessment of Educational Progress Improvement Act adopted as part of the Hawkins-Stafford Elementary and Secondary Education Amendments of 1988. That Act (PL 100 - 297), which passed following lengthy discussion, authorized voluntary state-by-state comparisons of NAEP assessment results on a trial basis, and mandated an independent study of the validity and effects of the pilot programs.

Less than two years later, prior to completion of the trial comparisons and the studies, NAGB is proposing a major expansion of NAEP (see NAGB's paper "Positions on the Future of the National Assessment"). The proposal includes: 1) full participation by the states in state-by-state comparisons, to be paid for by the federal government; 2) testing and comparing local districts and even schools, which is currently prohibited by law; and 3) more frequent testing. Last month, NAGB adopted a process for setting "achievement levels" that students in grades four, eight and twelve ought to attain on NAEP tests (see NAGB paper, "Setting Appropriate Achievement Levels").

While each of these initiatives raises problems that require serious attention, we are particularly concerned about the *combination* of setting achievement levels and expanding NAEP. Our specific concerns and recommendations include:

1) The proposal to expand NAEP was adopted before completion of the Congressionally-mandated studies or the pilot state-by-state comparisons.

Expansion of NAEP will inevitably affect our nation's education. Congress correctly planned a cautious, step-by-step process to gauge the value and effects of state comparisons before mandating their continuation or expansion. This evaluation should be completed before any further steps are taken to expand NAEP.

2) NAGB is proposing expansion of NAEP before the national debate on educational goals is resolved.

342 Broadway, Cambridge, Mass. 02139 (617) 864-4810 FAX (617) 497-2224

So far, the Bush Administration and the Governors have agreed on broad national goals, but have yet to decide how to implement them. Logically, the Administration, the Governors and Congress should all have roles in this debate as well as in determining the indicators used to measure progress toward the goals. But if measurement precedes goals clarification, the process of measuring becomes, by default, the process of defining. That would truly be putting the cart before the horse.

Deferring action on NAEP expansion until after the trial state comparisons and the legally required studies are completed will allow time for the national debate on attaining educational goals to reach resolution. Only then can NAEP play a constructive role in developing appropriate measurement tools and procedures.

3) It is reckless to consider lifting the ban on district-by-district or school-by-school comparisons without considering the consequences for curriculum and instruction.

No one yet knows the effects - and side-effects - even of state by state comparisons. Repeal of the ban on local comparisons requires much more information and public discussion. It should not be considered until after the results of all trial comparisons and the mandated studies have been fully analyzed and publicly discussed.

4) NAGB's achievement level setting process, when combined with comparisons, may create a *de facto* national curriculum.

The evidence is overwhelming that the more power attached to a test, the more control the test will have over curriculum and instruction. A national test with achievement goals and local comparisons will certainly become a powerful, perhaps controlling, influence on the curriculum.

The education goals enunciated by the Administration and the Governors do not attempt to mandate a national curriculum. In fact, there is widespread agreement that curriculum and instruction should not be determined from Washington. States and communities need flexibility in determining how to attain the broad goals. Yet NAGB's expansion proposals could preclude state and local initiatives.

5) NAGB's achievement level setting procedures for its tests are not appropriate for determining national achievement goals.

The process chosen by NAGB to set achievement levels on its tests relies on selecting items from existing NAEP exams that, in the view of committees of experts, should be answered correctly by students who have attained the levels of "basic," "proficient" or "advanced." This is not an appropriate method for determining national curricular goals and achievement levels because it allows one test to define the content area and what students should be able to do in that area. Such decisions should be made prior to and independently of any test. After curricular goals have been decided at the various levels, then assessments appropriate to the curriculum can be constructed and achievement levels set.

Moreover, as the recently-released report of the National Commission on Testing and Public Policy explains, the procedure of relying on committees of experts to examine items is

flawed even for the purpose of setting cut-off scores on tests. NAGB thus expects a limited technical procedure to be adequate for shaping a national curriculum.

6) By setting achievement goals based on what are predominantly multiple-choice tests, NAGB runs the risk of defining national educational goals in terms dictated by these narrow instruments.

In potentially shaping curriculum and instruction, NAEP tests will affect both content and methods of teaching. Multiple-choice testing necessarily focuses on factual recall and simple comparisons and observations. It does not lend itself to revealing whether students know how to do something - to write a persuasive essay, research an historical event, or grasp the meaning of a scientific development.

The narrowness of these instruments has been recognized by the Governors, among many others, and has led to widespread efforts to develop and implement other means of assessment. If multiple choice testing continues to predominate, NAEP will provide a continual obstacle to teaching and assessing the important things students need to learn how to do. It will help perpetuate a reduced definition of the content to be studied and an entirely incorrect view of how students learn.

7) NAGB proposes to vastly increase the amount of its testing to include "at least three subjects each year."

The current NAEP authorization establishes a two-year testing cycle and a minimum frequency for testing various subjects. Only math and reading are to be tested every two years; other subjects are scheduled at four- or six-year intervals. Though its futures paper deferred discussion of the "exact configuration" of the new testing cycles, NAGB called for "testing at least three subjects each year," at least six tests every two years. NAGB's claims this acceleration is necessary "to provide timely and sufficient data" and to "replace the Education Department's annual 'wall chart' which relies on SAT and ACT scores."

Again, major changes in NAEP such as expanding the extent and frequency of testing should not be undertaken prior to completion and analysis of the 1992 testing and the mandated studies. In fact, such expansion is not at all necessary. Because educational systems and achievement cannot change rapidly, yearly aggregated data will not provide meaningful information about important educational changes. Less frequent information should be quite sufficient.

While virtually everyone, including Secretary Cavazos, agrees on the inadequacy of the current "wall charts," the mere existence of the charts is an insufficient justification for vastly increasing a national testing program. To be sure, annual one-point changes in average SAT scores or two-tenths of a point changes on the ACT in the "wall charts" are meaningless. But substituting minute changes in NAEP scores would not be an improvement. It could, however, produce public frustration and thereby jeopardize public support for educational reform. Maintaining NAEP's current, authorized schedule will provide as much useful information at less cost in dollars and, ultimately, in public credibility.

8) NAGB is moving too slowly in revising NAEP exams to rely less on multiple-choice questions and to develop other means of assessment which better measure the full range of knowledge and skills.

While NAGB claims that about 20% of this year's NAEP math items were open-ended, Paul LeMahieu, Pittsburgh's Director of Testing, informed the National Association of Test Directors that less than 5% were really open-ended items. The rest were multiple-choice questions with the answer options deleted. Like multiple-choice items, such questions are not very useful in measuring student abilities to use math to solve real-world problems.

Instead of expanding the use of outdated, multiple-choice tests, NAEP should become a leader in the national effort to develop improved forms of assessment that provide more information and do not endanger but rather enrich the curriculum. NAEP should work with the states, a number of which already have performance-based assessment projects under development, to produce and evaluate such assessments.

9) NAEP expansion will absorb an ever larger share of federal research and information dollars, but the results may not be worth the money.

The NAEP Improvement Act authorized $9,500,000 for fiscal year 1989 for NAEP. For FY 1990, NAEP received $17,084,000. Even with this increased amount, the Education Department deferred the NAEP validity study, a national assessment of adult illiteracy and work on the National Education Longitudinal Study. For FY 1991, NAGB has requested $18,866,000, an increase of more than 10% over FY90 and nearly double the authorization for FY89. NAGB receives up to 10% of NAEP funds for administrative purposes and reportedly seeks to receive up to 15%. Estimates of the cost of NAEP if expanded are $100 million annually, a more than five-fold increase over current expenditures and an amount two and one-half times the funding for the National Center for Educational Statistics (NCES).

Will the results be worth the additional money? Yearly testing will not increase anyone's knowledge of the effects of educational reform efforts. Further state and local comparisons may not tell us more than we already know about how well the states and localities perform on standardized tests. In a period of continuing fiscal restraint, money used for more extensive testing could be better used to improve the *quality* of NAEP assessments or for other needed research rather than for redundant and potentially dangerous increases in testing.

10) The relationship among NAEP, NAGB and NCES must be clarified.

The current debates over the future of NAEP have raised questions about the appropriateness of an independent body wielding the power that NAGB could assert over our nation's education. A key issue is whether such a body is adequately accountable to Congress, the Administration and the public.

Since accountability is, in part, asserted by control over funding, NAGB's budget should be separated from NAEP's. So long as NAGB obtains a percentage of a (potentially rapidly-expanding) NAEP budget, there is no way for elected officials to adequately exert oversight. The role of NAGB in relation to NCES, the Department of Education or any other bodies created to oversee progress toward national goals should be carefully considered by the appropriate House and Senate committees and the Administration before NAEP is expanded.

In sum, NAGB's plans to rapidly expand NAEP without adequate consideration of the effects of the expansion or the proper role of assessment in educational reform are dangerous. Neither Congress nor the administration should allow them to proceed without careful review and consideration. Similarly, the Governors should not support the use of NAEP for measuring progress toward national goals without first clarifying the goals and the role of assessment in achieving them and then determining the details of measurement. Specifically:

- NAEP should not be expanded to allow more frequent or extensive testing or more detailed comparisons at least until completion of the trial assessments of 1990 and 1992 and the independent evaluation mandated in the Act. Then, Congress, the Administration and the Governors must weigh carefully the potentially harmful effects of more extensive testing and comparisons and ascertain that the dangers do not outweigh any possible benefits. In any event, expansion of NAEP must be subsequent and subordinate to the establishment of national goals and not allowed to dictate a national curriculum.

- NAEP should be directed to spend a significant portion of its budget on developing and piloting performance-based assessments (including tests and portfolios). Such research and development should be planned carefully to coordinate with state projects such as those underway in California, Connecticut and Vermont, to develop performance-based assessments, as well as projects undertaken by local education authorities or other governmental or private bodies.

- Congress and the Administration should consider separating NAGB funding from NAEP funding and carefully consider the future role of NAGB in relation to other agencies and bodies.

We appreciate your attention to these most important issues and look forward to working with you in the effort to achieve genuine and lasting reforms in the quality of public education.

Please feel free to call any of us if you have any questions or need further information.

List of Signers is Attached

LIST OF SIGNERS

Ronald J. Abate, College of Education, Cleveland State University*
Advocates for Children of New York, Norman Rollins, Executive Director
American Association of Colleges for Teacher Education, David G. Imig, Executive Director
American Association of School Administrators
American Reading Council, Julia Palmer, President
Angelo N. Ancheta, Asian Law Alliance*
APPLE Corps., Inc., Alfred E. McWilliams, President; MaryAnne F. Gaunt, Executive Director
Association for Supervision and Curriculum Development, Gordon Cawelti, Executive Director
Bank Street College of Education, Joseph Shenker, President
Rims Barber, Mississippi Human Services Agenda
Leonard Beckum, Dean, College of Education, City College of New York
Association for Women in Science, Stephanie J. Bird, President
Gerald W. Bracey, Director of Research and Evaluation, Cherry Creek, Colorado
Diana Caballero, Puerto Rican/Latino Education Roundtable, New York City*
Center for Law and Social Policy, Alan W. Houseman, Director
Center for Women Policy Studies, Leslie R. Wolfe, Executive Director
Harold E. Dent, Vice-President, Psychological and Human Resources Consultants, Inc.*
Educational Law Center, Inc. Marilyn Morheuser, Executive Director
Pablo Eisenberg, President, Center for Community Change*
Federation of Organizations for Professional Women
Foxfire Fund, Inc. Eliot Wigginton, President
Fund for the Feminist Majority, Eleanor Smeal, President
Howard Gardner, Project Zero, Harvard University*
Genesee Valley Developmental Learning Group, New York
Leslie A. Hart, Brain-Compatible Education Associates
La Donna Harris, Americans for Indian Opportunity*
Asa Hilliard, Professor of Education, Georgia State University*
Holt Associates/*Growing Without Schooling*
Institute for Learning and Teaching, Wayne Jennings, Director
International Reading Association
KEY-Kids, Education and You, (Beth Bradley, Jenny Coston, Leslie Floyd, Sue Long)
A. Gay Kingman, National Congress of American Indians*
Nancy K. Klein, College of Education, Cleveland State University*
Massachusetts Advocacy Center, Stephen R. Bing, Executive Director
Matsushita Foundation, Inc., Dr. Sophie Sa, Executive Director
Deborah Meier, Principal, Central Park East Secondary School, New York City
Sara E. Melendez, Vice-Provost, University of Bridgeport*
Susan Metz, Prospect Heights High School, Brooklyn, NY*
Mexican American Women's National Assocaition
National Association for the Advancement of Colored People, Beverly Cole, Education Director
National Association of Secondary School Principals
National Center for Fair & Open Testing (FairTest), Cinthia Schuman, Executive Director
National Coalition of Advocates for Students
National Coalition of Title I/Chapter I Parents, Robert Witherspoon, Director
National Council for the Social Studies
National Council of Teachers of English, Charles Suhor, Executive Director
National Education Association, Keith Geiger, President
National Indian Youth Council, Inc., Cheryl J. Mann, Executive Director
National Organization for Women--New York City Chapter

National Parent Teacher Association
National Women's Law Center
National Women's Political Caucus
Fred M. Newmann, National Center on Effective Secondary Schools--University of Wisconsin*
New York Public Interest Research Group
NOW Legal Defense and Education Fund, Helen Neuborae, Executive Director
Organization of Chinese American Women, Faith Lee Breen, Chair, Board of Directors
Vito Perrone, Harvard University--Graduate School of Education*
Project Equality, Inc., Kansas City, MO
Puerto Rican Legal Defense and Education Fund, Inc., Ruben Franco, President/General Counsel
Rochester (NY) Teachers Association (AFT), Adam Urbanski, President
Lori Rubenstein, Partnership for Democracy*
William V. Schipper, Executive Director, Natl. Assn. of State Directors of Special Education*
Donald H. Smith, Chairman, Dept. of Education, Baruch College, City University of New York*
Hilton Smith, Coordinator, Foxfire Teacher Outreach Programs*
Southern Association on Children Under Six, Cathy Grace, Executive Director
Southern Christian Leadership Conference, Joseph Lowery, President
Southern Regional Council, Inc, Steve Suitts, Executive Director
Judy I. Stahlman, Cleveland State University*
Adria Steinberg, Harvard Education Letter*
Gail E. Thomas, Texas A & M University*
United States Student Association, Julianne Marley, President
Dorothy J. Watson, President, Whole Language Umbrella*
Paul Weckstein, Center for Law and Education*
Grant Wiggins, Center on Learning, Assessment and School Structures (CLASS)
Arthur E. Wise, Rand Corporation*
Women's Research and Education Institute, Betty Parsons Dooley, Executive Director

***Organization for identification purposes only**